Warning Concerning Copyright Restrictions

The Copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyright material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction not be "used for any purposes other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

CHAPTER 2

SHORT-TERM MEMORY

Alan Baddeley

In 1887, John Jacobs, a schoolmaster in London, wanted to assess the abilities of his students. He devised an apparently simple test in which the student heard a sequence of digits, like a telephone number and repeated them back. The measure used was **digit span**, the longest sequence that could be repeated back without error (Jacobs, 1887). Digit span is still included in the most widely used intelligence test, the Wechsler Adult Intelligence Scale (WAIS). In this basic version, span does not correlate very highly with general intelligence but, as we will see, a somewhat more complex version, **working memory span**, does an excellent job of predicting a wide range of cognitive skills, including performance on the reasoning tasks often used to assess intelligence.

The digit span test is typically referred to as reflecting *short-term memory* (STM), and the more complex task as *working memory span*. The terms short-term memory (STM) and working memory (WM) seem often to be used interchangeably, so is there a difference?

SHORT-TERM AND WORKING MEMORY: WHAT'S THE DIFFERENCE?

The term "short-term memory" is a rather slippery one. To the general public, it refers to remembering things over a few hours or days, the sort of capacity that becomes poorer as we get older and is dramatically impaired in patients with Alzheimer's disease. To psychologists, however, these are long-term memory (LTM) problems. Remembering over a few minutes, hours, or a few years all seems to depend on the same long-term memory system.

We will use the term *short-term memory* (STM) to refer to performance on a particular type of task, one involving the simple retention of small amounts of information, tested either immediately or after a short delay. The memory system or systems responsible for STM form part of the *working memory* system. "Working memory" is the term we will use for a system that not only temporarily stores information but also manipulates it so as to allow people to perform such complex activities as reasoning, learning, and comprehension. Before going on to discuss working memory in the next chapter, we will examine the simpler concept of STM, the capacity to store small amounts of information over brief intervals, beginning with the digit span task devised by Jacobs.

We will use the term STM to describe the process of storing small amounts of

KEY TERMS

Digit span: Maximum number of sequentially presented digits that can reliably be recalled in the correct order.
Working memory span: Term applied to a range of complex memory span tasks in which simultaneous storage and processing is required.

information over a brief interval. As such, the term itself is theory-free, although there are of course theories as to how this is achieved.

In contrast to our use of STM simply to describe an experimental situation, the term *working memory* is based on a theoretical assumption, namely that tasks such as reasoning and learning depend on a system that is capable of temporarily holding and manipulating information, a system that has evolved as a mental work space. A number of different theoretical approaches to working memory have developed, some influenced strongly by the study of attention (e.g. Cowan, 2001), some on studies of individual differences in performance on complex tasks (e.g. Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000; Engle & Kane, 2004), and others driven by neurophysiological considerations (Goldman-Rakic, 1996). All, however, assume that WM provides a temporary workspace that is necessary for performing complex cognitive activities.

The approach used in the next two chapters reflects a multicomponent account of WM (Baddeley & Hitch, 1974) that was strongly influenced by the experimental and neuropsychological studies of human memory that form the core of the present book. It has proved durable and widely applicable, but should be seen as complementary to a range of other approaches rather than as *the* theory of working memory (Miyake & Shah, 1999a). The multicomponent model has relied heavily on studies of STM and has, as its most thoroughly explored component, the *phonological loop*, which provides a theoretical account of verbal STM. The present chapter is concerned with the major findings from studies of verbal and visual STM. It uses the phonological loop model to tie them together but also refers to other competing theories of STM. This leads on to Chapter 3, which goes beyond the study of STM, with its emphasis on short-term storage, to consider the question of why such temporary storage is needed and how it is used as part of a much broader concept, that of working memory.

MEMORY SPAN

Before proceeding, test yourself using Box 2.1.

If your digit span is rather lower than you might hope, don't worry; in this simple form, as we shall see later, it depends on a small but useful aspect of our memory system, not on general intelligence. Digit span is a classic short-term memory task, in that it involves holding a small amount of material for a short period of time.

The most obvious fact about digit span is that it is limited to about six or seven digits for most people, although some people can manage up to ten or more, whereas others have difficulty recalling more than four or five. What sets this limit and why does it vary between one person and the next?

Memory span measures require two things: (1) remembering what the *items* are; and (2) remembering the *order* in which they were presented. In the case of the digits one to nine, we already know the items very well,

Box 2.1 Digit span test

Read each sequence as if it were a telephone number, then close your eyes and try to repeat it back. Start with the four digit numbers, and continue until you fail on both sequences at a given length. Your span is one digit less than this.

9 7 5 4
3 8 2 5
9 4 3 1 8
6 8 2 5 9
9 1 3 8 2 5
6 4 8 3 7 1
7 9 5 8 4 2 3
5 3 1 6 8 4 2
8 6 9 5 1 3 7 2
5 1 7 3 9 8 2 6
7 1 9 3 8 4 2 6 1
1 6 3 8 7 4 9 5 2
9 1 5 2 4 3 8 1 6 2
7 1 5 4 8 5 6 1 9 3

so the test becomes principally one of memory for order. If, however, I were to present you with a sequence of digits in an unfamiliar language, Finnish for example, your span would be very much less. You would of course have much more to remember, as you would need to recall the order of the sounds comprising the Finnish digits, as well the order of the digits.

Suppose I were to use words, but not digits. Would that matter? Provided I used the same words repeatedly, you would soon become familiar with the set, and would do reasonably well. However, if I were to use a different set of words on each trial it would become somewhat harder as you would again need to remember both what the items were and their order, although this would be easier than for the unfamiliar Finnish digits.

How is order remembered? Not an easy question to answer. One might think that each number is associated or linked to the next, which is turn is linked to the next, a process known as *chaining*. The problem here is that if one link of the chain breaks, because one item is forgotten, then performance on all later items should collapse. In fact, although there is an increase in errors following a mistake, the subsequent decline is by no means as dramatic as chaining would predict. As we shall see, the simple task of explaining how the brain stores sequences remains controversial, not only for remembering words and numbers but also because learning and reproducing sequences of actions is essential for a wide range of activities, from simple reaching and grasping to the performance of a skilled athlete. Suppose we move from numbers to letters. Test yourself on the next sequence by reading each letter out loud, then close your eyes and try to repeat the letters in the order they are written.

C T A I I L T C S F R O

Now try the next sequence.

F R A C T O L I S T I C

I assume you found the second sequence easier, even though it used exactly the same letters as the first. The reason is that the order of the letters in the second sequence allowed you to break it up into pronounceable word-like subgroups or *chunks*. In a classic paper, George Miller (1956) suggested that memory capacity is limited not by the number of *items* to be recalled, but by the number of *chunks*. The first sequence comprised twelve apparently unrelated letters, making it hard to reduce the number of chunks much below twelve, whereas the second could be pronounced as a string of four syllables that, together, made a sequence that, although meaningless, could plausibly be an English word.

Chunking in this case depends on letter sequences that are consistent with long-term language habits. Grouping can also be induced by the rhythm with which a sequence of items is presented. Suppose I were to read out nine digits. If I interposed a slightly longer pause between items three and four and items six and seven, recall would be significantly improved. Hence 791-684-352 is easier than 791684352. Pauses in other locations can also be helpful, but grouping in threes seems to be best (Wickelgren, 1964; Ryan, 1969). It seems likely that the memory system is taking advantage of cues from prosody, the natural rhythms that occur in speech and that make its meaning clearer by separating into coherent phrases the continuous sequence of sounds that make up the normal speech stream.

Although remembering strings of numbers was probably of little interest to Mr Jacobs' students, it has in recent years become much more critical because of the increasing use in our culture of digit and letter sequences, initially as telephone numbers, then as zip codes, and subsequently as PINs and passwords. In the early 1960s, Dr R Conrad was tasked by the British Post and Telecommunications Service to investigate the relative advantages and disadvantages of codes based on letters and

KEY TERM

Chunking: The process of combining a number of items into a single chunk typically on the basis of long-term memory.

numbers. One of his experiments involved visually presenting strings of consonants for immediate recall. He noticed an interesting pattern in his results, namely that, despite being presented visually, errors were likely to be similar in *sound* to the item they replaced, hence *P* was more likely to be misremembered as *V* than the more visually similar letter *R* (Conrad, 1964). Conrad and Hull (1964) went on to investigate this effect further, demonstrating that memory for sequences of consonants is substantially poorer when they are similar in sound (e.g. *C V D P G T* versus *K R X L Y F*). Conrad interpreted his results in terms of a short-term memory store that relies on an acoustic code, which fades rapidly, resulting in forgetting. This was assumed to be particularly disruptive of recall of the acoustically similar letters as they had fewer distinguishing features, making each item more likely to be confused with adjacent items, resulting in errors in order of recall (e.g. *PTCVB* recalled as *PTVCB*).

TWO KINDS OF MEMORY?

Have you ever had the experience of walking from one room to another to pick something up, and by the time you arrive forgetting why you set off? If so, it is probably because you are thinking about something else, so that the original intention just slipped your mind. Lloyd and Margaret Peterson (1959), from the University of Indiana, developed a technique that could be seen as an analog of this experience. Participants were given an item to remember, a consonant triplet such as *XRQ*, and then distracted by being required to count backwards in threes from a given number (e.g. *371*: 368, 365, 362, etc.). After varying numbers of seconds of counting, they were asked to recall the triplet. Using a similar design, Murdock (1960) showed an equivalent effect for triplets of words (Figure 2.1).

How would you explain this? One possibility is that the numbers are interfering with memory of the letters. As we shall see in Chapter 9, there is abundant evidence that learning can be disrupted by subsequent activity. However, earlier research had shown that such interference depends on the *similarity* between the remembered and interfering material, and that for LTM at least, numbers do not interfere with letters (McGeogh & MacDonald, 1931). The Petersons therefore suggested their results reflected the rapid fading of a short-term memory trace. This interpretation was consistent with the

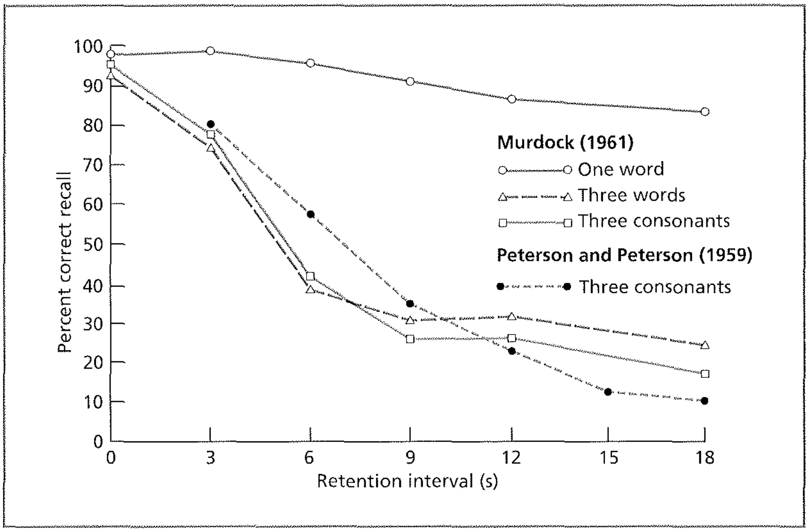
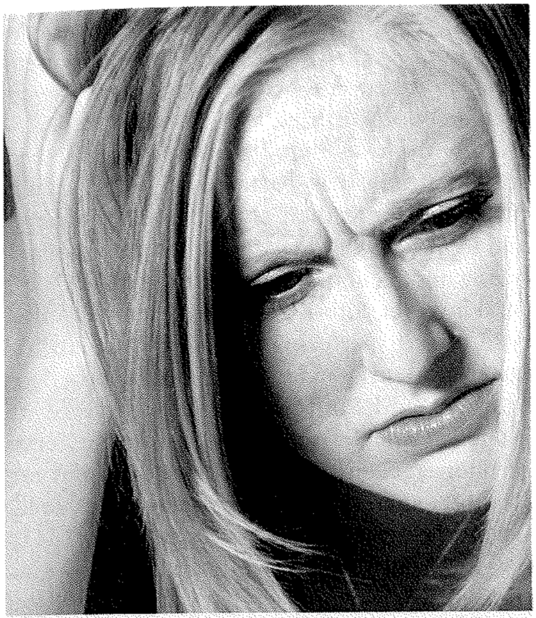


Figure 2.1 Short-term retention of consonant trigrams (Peterson & Peterson, 1959) and of one-word and three-word sequences (Murdock, 1961). From Melton (1963).

conclusion drawn by John Brown (1958) in England, based on a similar demonstration of short-term forgetting. It was however directly at odds with the widely held view at that time of memory as a unitary system, in which forgetting results from interference.


Forgetting, as a result of interference, depends on the similarity between the remembered and interfering material. We are more likely to forget our original intention when distracted by another similar thought or task.

A simple trace decay interpretation was, however, challenged by Keppel and Underwood (1962), who showed that rapid forgetting was something that built up over the first four or five trials of the experiment. The very first triplet presented showed little or no forgetting. They argued that their results suggested the Peterson effect was attributable to interference from *earlier* triplets, which were, of course, similar in type to the remembered items. This hypothesis can be tested by changing the nature of the remembered items from one trial to the next.

As mentioned earlier, interference depends on similarity. Hence, if the previous item is not similar it should not cause forgetting. This was tested in a further experiment using word triplets in which each item in the given triplet came from the same semantic category: three birds, for example, or three colors. After five trials, all based on different examples of the same category (birds), the category was changed (e.g. to colors). As Figure 2.2 shows, performance declined steadily over the first five trials and recovered once the new category was introduced, only to decline once more until the category was changed once again (Loess, 1968).

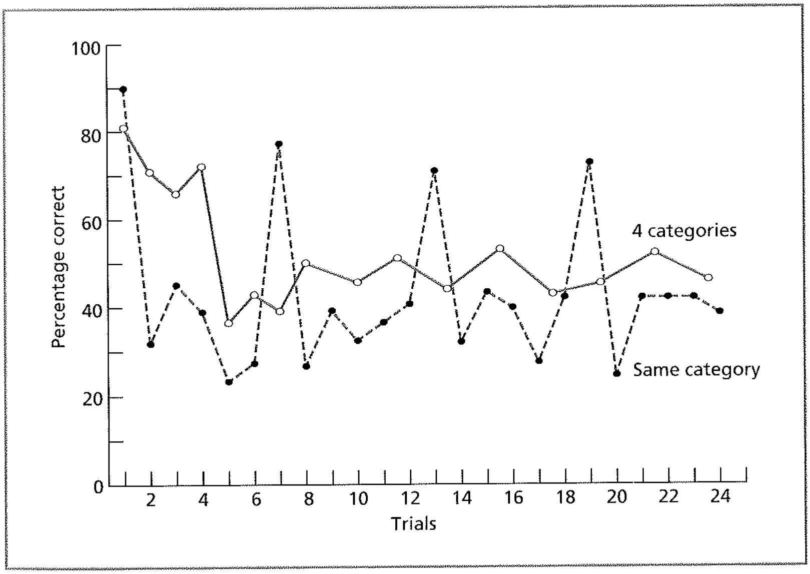


Figure 2.2 The release from the proactive interference effect. Participants were required to recall sets of four items from a single taxonomic category. After six sets, the category changed, resulting in enhanced performance on this set. Sets comprising items from all four categories show a steady decline and no release from proactive interference. From Loess (1968). Copyright © Elsevier. Reproduced with permission.

The simplest way of thinking about the **Peterson task** is in terms of the problem the participant has to solve, namely that of reproducing the *last* of a number of sequences that have been presented. The longer the delay before recall, the more difficult it is to distinguish which was the last, and which occurred slightly earlier. Considerable controversy ensued as to whether this could or could not fit into the existing stimulus–response interference theory (see Baddeley, 1993, pp. 31–37) however, there is broad agreement that the Peterson task demands on explanation in terms of retrieval processes that go beyond the initial hypothesis of simple trace decay.

Before going on to discuss further the question of trace decay versus interference, we should consider a second experimental paradigm that became both popular and theoretically controversial during the 1960s, namely **free recall.**

Free recall

In this task, participants are simply given lists of items, which they must then recall in any order they wish. Figure 2.3 shows the results of an experiment by Postman and Phillips (1965) in which 10, 20, or 30 words were presented, to be recalled immediately or after a 15-second filled delay. Their results illustrate a number of features characteristic of free recall including: (1) that the likelihood of recalling an *individual* item is less for longer lists, although the *total* number of items recalled is likely to increase; (2) that all lists showed a tendency for the first few items to be somewhat better recalled, the so-called **primacy effect**; (3) that, regardless of list length, if recall is immediate then the last few items are very well recalled, the **recency effect**; and (4) that this effect is eliminated by a brief delay filled by some activity such as counting.

How should we interpret this pattern of results? Evidence indicates that performance on the earlier items depends principally on LTM, with the primacy effect probably being due at least in part to the tendency to rehearse the first few items as they come in, sometimes perhaps continuing to rehearse these items throughout the list (Rundus, 1971; Hockey, 1973; Tam & Ward, 2000). A whole range of variables is known to affect LTM influence performance on the early and middle part of the list (Glanzer, 1972). These include: (1) *presentation rate*: slower is better; (2) *word frequency*: familiar words are easier; (3) *imageability of the words* words that are visualizable are better; (4) *age of the participant*, young adults remember more than children or the elderly; and (5) *physiological state*: drugs such as marihuana and alcohol impair performance.

However, although these factors all influence overall performance, none of them has an equivalent impact on the recency effect. Postman and Phillips themselves offered an interpretation of their own results in terms of interference theory, although this does not readily explain why these factors influence overall retention but not recency.

The most popular interpretation during the 1960s was to assume that the recency effect reflected a temporary short-term store, which had different characteristics from the long-term store responsible for performance on earlier items (Glanzer, 1972). However, the assumption that recency simply reflects the output of a short-term store was subsequently challenged by the demonstration that recency effects can occur under conditions in which the short-term trace ought to have been disrupted. In one

KEY TERMS

The Peterson task: Short-term forgetting task in which a small amount of material is tested after a brief delay filled by a rehearsal-preventing task.

Free recall: A method whereby participants are presented with a sequence of items, which they are subsequently required to recall in any order they wish.

Primacy effect: A tendency for the first few items in a sequence to be better recalled than most of the following items.

Recency effect: A tendency for the last few items in a list to be well recalled.

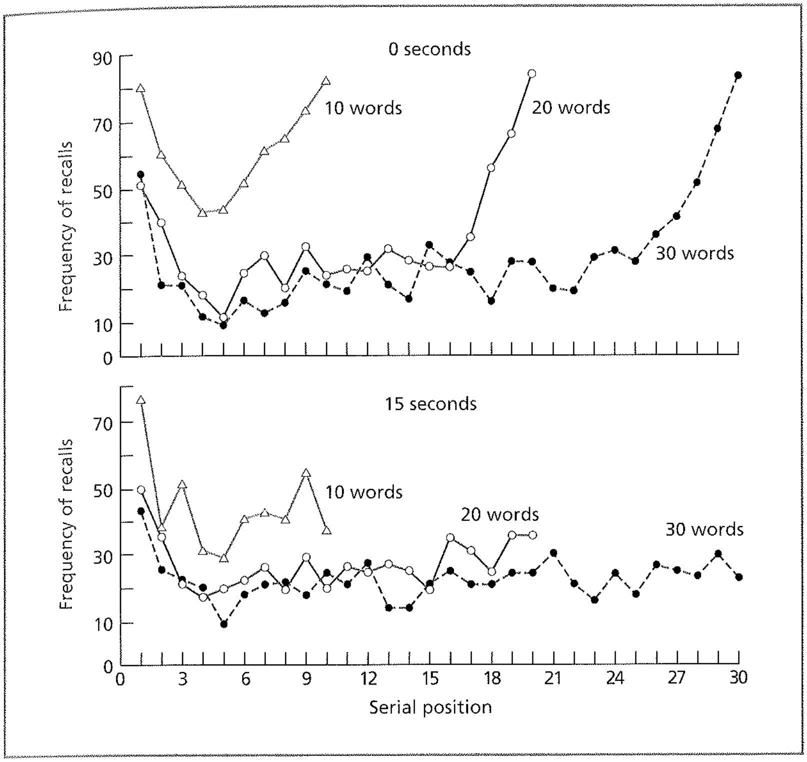


Figure 2.3 Serial position curves for lists of 10, 20, or 30 words recalled immediately or after a 15-second delay. Note that for each list length the last few items are very well recalled on immediate test—the recency effect—but not after the delay. From Postman and Phillips (1965). Copyright © Psychology Press.

study, Bjork and Whitten (1974) required their subjects to recall sequences of words presented under three conditions. The baseline condition involved presenting a list of words for immediate free recall. As expected, this resulted in a clear recency effect. In a second condition, the gap between presentation and recall was filled by a 20-second backward counting task, which—as expected—removed the recency effect. In a third, crucial condition, 20 seconds of backward counting was interposed between each of the words presented, as well as between the end of the list and recall. Under these conditions, a recency effect re-emerged.

Recency effects have also been demonstrated over much longer intervals. In one study, for example, Baddeley and Hitch (1977) tested the capacity of rugby players to recall which teams they had played that season; their recall showed a clear recency effect. As not all the players had played in all the games, it proved possible to assess whether forgetting was more reflective of the amount of time elapsed, or of the number of intervening games. Number of games proved to be the better predictor, suggesting that a simple decay hypothesis would not provide a good account of these findings. Similar **long-term recency** effects have been found in remembering parking location (Pinto & Baddeley, 1991), although—sadly—I can report that as I get older, even recency does not always prevent the need for an embarrassed wander around the grocery store parking lot.

The fact that recency effects are found across such a wide range of situations, with some cases being disrupted by a few seconds of unrelated activity such as counting whereas others persist over months, suggests that the recency effect is not limited to any single

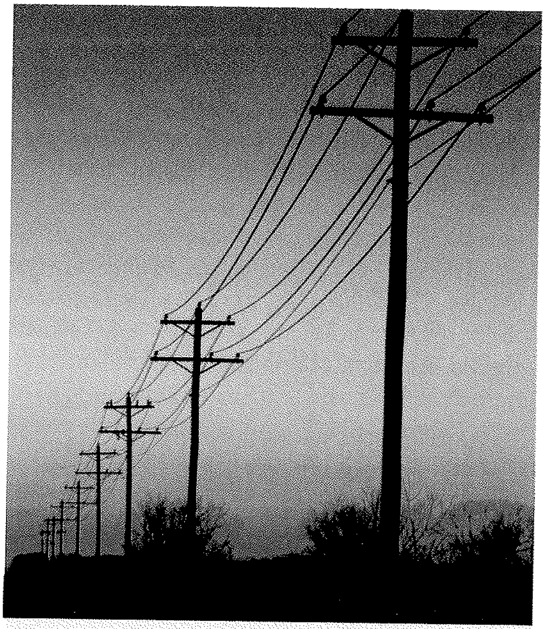
KEY TERM

Long-term recency: A tendency for the last few items to be well recalled under conditions of long-term memory.

type of memory system but instead reflects a specific retrieval strategy that takes advantage of the fact that the most recent events are the most readily available to recall.

When was the last party you attended? Which was the party before that? And the one before? I suspect recalling your most recent party was the easiest, although it was perhaps not the best party.

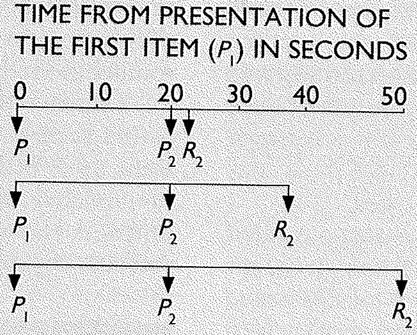
The greater accessibility of the most recent experience of a given type could serve the highly important role of orienting yourself in space and time. When traveling and staying in a new hotel, how do you know where you are when you wake up? And when you leave the hotel for dinner, how do you remember your current room number and don't recall instead the number from last night or the night before that?



Crowder's (1976) analogy likened the task of retrieval from a free recall list to that of discriminating between a string of telephone posts; the further away the post is from the observer, the more difficult it becomes to distinguish it from its neighbor.

As in the case of the Peterson effect, the most plausible interpretation of recency seems to be in terms of retrieval. Crowder (1976) likens the task of retrieving items from a free-recall list to that of discriminating telephone posts located at regular intervals along a railway track. The nearest post will be readily distinguishable from the next, while as the posts recede into the distance, the problem of separating one from the other becomes increasingly hard (Box 2.2). Both this and the Peterson effect can be seen in terms of a

Box 2.2 Recency



Delay (sec) at R_2		Discrimination Ratio
P_1	P_2	(P_1 : P_2)
22	2	11 : 1
30	10	3 : 1
50	30	1.67:1

P_1 = Presentation of item 1
P_2 = Presentation of item 2
R_2 = Recal of item 2

discrimination ratio, based on the temporal distance between the item being retrieved and its principal competitor. On immediate recall, the most recent item has a considerable advantage, but with increasing delay, discriminating the last item from the one before becomes less and less easy (Glenberg, Bradley, Stevenson, Kraus, Tkachuk, Gretz et al., 1980; Baddeley & Hitch, 1977, 1993; Brown, Neath, & Chater et al., 2007).

MODELS OF VERBAL SHORT-TERM MEMORY

By the late 1960s, the evidence seemed to be swinging firmly in the direction of abandoning the attempt to explain STM in terms of a unitary system, in favor of an explanation involving a number of interacting systems, one of which was closely identified with the extensive evidence accumulated from verbal STM. I will use one out of a range of models of verbal STM, the concept of a *phonological loop*, to tie together the rich body of research that continues to develop in this area, before going on to give a brief account of alternative theories.

The phonological loop

The concept of a **phonological loop** forms part of the multicomponent working memory model proposed by Baddeley and Hitch (1974). The phonological loop is assumed to have two subcomponents, a short-term store and an articulatory rehearsal process. The store is assumed to be limited in capacity, with items registered as memory traces, which decay within a few seconds. However, the traces can be refreshed by subvocal rehearsal, saying the items to yourself, which depends on a vocal or subvocal articulatory process.

Consider the case of digit span. Why is it limited to six or seven items? If there are few digits in the sequence, then you can say them all in less time than it takes for the first digit to fade away. As the number of items increases, total time to rehearse them all will be greater, and hence the chance of items fading before they are refreshed will increase, hence setting a limit to memory span. The loop model is able to account for the following prominent features of verbal STM:

The phonological similarity effect

A major signature of the store is the **phonological similarity effect**, Conrad's (1964) demonstration that letter span is reduced for similar sounding items. Thus, remembering a sequence of five dissimilar words (e.g. *pit*, *day*, *cow*, *pen*, *hot*) is relatively easy, whereas remembering those that are phonologically similar (e.g. *cat*, *map*, *man*, *cap*, *mad*) is much harder. Note, however, that this is not a general effect of similarity, as sequences that are similar in meaning (e.g. *huge*, *wide*, *big*, *long*, *tall*) are only slightly more difficult than dissimilar sequences (e.g. *old*, *wet*, *thin*, *hot*, *late*) (Figure 2.4).

One final point to this story is that the phonological similarity effect disappears if the lists are increased in length and participants are allowed several learning trials. Under these circumstances, similarity of meaning becomes much more important (Baddeley, 1966b). This does not mean that phonological coding is limited to STM, as without phonological LTM we could never learn to pronounce new words. It does however mean that LTM gains more from relying on meaning than on sound. We return to this point in Chapters 4 and 5.

The phonological similarity effect is assumed to occur when information is read

KEY TERMS

Phonological loop: Term applied by Baddeley and Hitch to the component of their model responsible for the temporary storage of speech-like information.

Phonological similarity effect: A tendency for immediate serial recall of verbal material to be reduced, when the items are similar in sound.

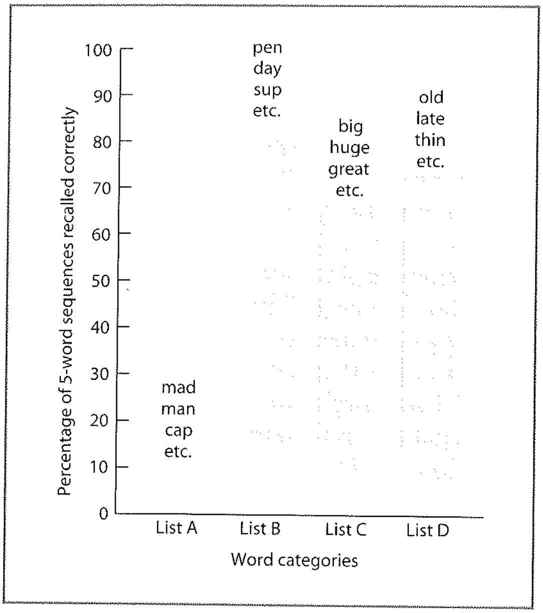


Figure 2.4 The effect of phonological and semantic similarity on immediate serial recall of five-word sequences. Phonological similarity leads to poor immediate recall whereas similarity of meaning has little effect. From Baddeley (1966a). Copyright © Psychology Press.

out from the short-term memory trace; similar items have fewer distinguishing features, and hence are likely to be confused. Auditory speech is assumed to feed directly into the phonological store. Visually presented items can also be fed into the store if they are nameable, such as digits, letters, or nameable objects, through a process of vocal or subvocal articulation, whereby you say the items to yourself.

The subvocal rehearsal system will be blocked if you are required to repeatedly say something unrelated such as the word "the," an activity known as **articulatory suppression.** Saying "the" means that you are not able to refresh the memory trace by subvocally pronouncing the remembered material. It also prevents you from subvocally naming visually presented items, such as letters, which prevents them from being registered in the phonological store. For that reason, it does not matter whether items are phonologically similar or not, when they are presented visually and accompanied by articulatory suppression. Both similar and dissimilar items will be retained at a lower but equivalent level.

However, it is important to note that even when suppressing, people can still remember up to four or five visually presented digits. This suggests that although the phonological loop typically plays an important role in digit span, it is not its only basis. We will return to this point later. With auditory presentation, the words gain direct access to the phonological store despite articulatory suppression, and a similarity effect occurs.

The word length effect

As we saw, most people can remember sequences of five dissimilar one-syllable words relatively easily. As word length increases, performance drops from around 90% for five monosyllables to about 50% for lists of five-syllable words. As word length is increased, the time taken to speak the words also increases (Figure 2.5). This relation between recall and the rate of articulation can be summarized by the statement that people can remember about as many words as they can say in 2 seconds (Baddeley, Thomson & Buchanan, 1975).

Interpreting these findings within the phonological loop model is relatively simple. Rehearsal takes place in real time, as does trace decay, with the result that longer words, taking longer to say, allow more decay to occur. Baddeley et al. (1975) attributed the **word length effect** to forgetting during rehearsal. However, Cowan, Day, Saults, Keller, Johnson, and Flores (1992) demonstrated that word

KEY TERMS

Articulatory suppression: A technique for disrupting verbal rehearsal by requiring participants to continuously repeat a spoken item.

Word length effect: A tendency for verbal memory span to decrease when longer words are used.

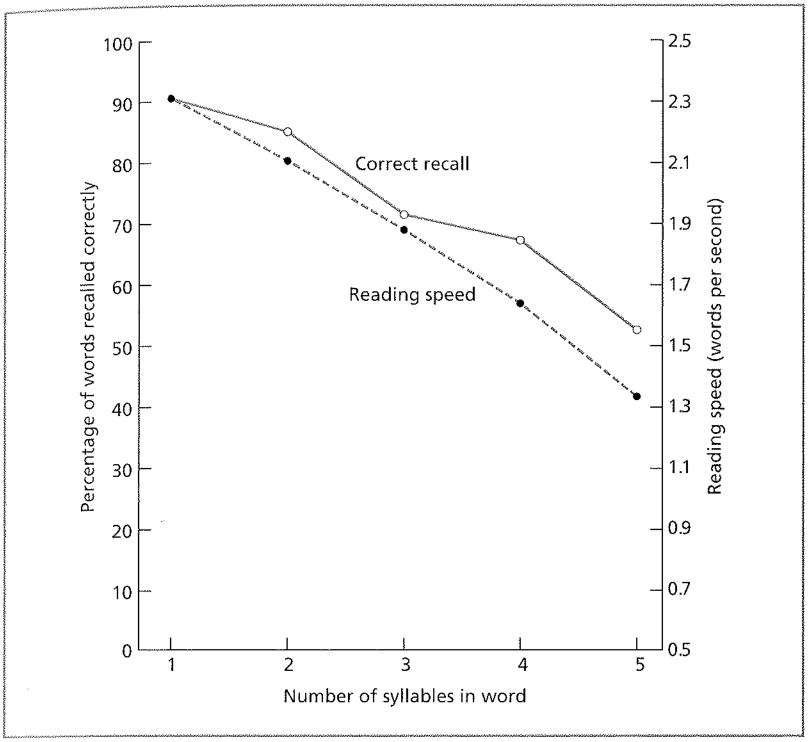


Figure 2.5 The relationship between word length, reading rate, and recall. Long words take longer to rehearse and also produce lower memory spans. From Baddeley, Thomson, and Buchanan (1975). Copyright © Elsevier. Reproduced with permission.

length also caused forgetting due to the fact that longer words take longer to recall, and Baddeley, Chincotta, Stafford, and Turk (2002) showed that both effects occur. Articulatory suppression, by requiring participants to utter an irrelevant sound, stops the process of verbal rehearsal. This abolishes the word length effect. Unlike the phonological similarity effect, this disruption occurs regardless of whether presentation is auditory or visual; if articulation is blocked, neither method of presentation will allow verbal rehearsal, and hence the word length effect is abolished (Baddeley et al., 1975).

The word length effect is extremely robust but its interpretation remains controversial. An alternative to the Baddeley et al. (1975) suggestion that the effect was based on trace decay, with longer words taking more time to rehearse and recall, is the proposal that longer words are more complex and this leads to more interference (e.g. Caplan, Rochon, & Waters, 1992). A third interpretation suggests that long words, having more components to be remembered, are more vulnerable to fragmentation and forgetting (e.g. Neath & Nairne, 1995), although this interpretation has now been abandoned by its earlier proponents (Hulme, Neath, Stuart, Shostak, Suprenant, & Brown, 2006). However, the controversy is still far from resolved (see Mueller, Seymour, Kieras, & Meyer, 2003; Hulme et al., 2006; Baddeley, 2007; Lewandowski & Oberauer, in press, for further discussion of this complex issue).

Happily, although it is an important issue, the question of whether the greater degree of forgetting observed with long words is due to time-based trace decay, or interference from subsequent syllables is not crucial to the general concept of a phonological loop operating as part of a multicomponent working memory model. This is fortunate, as the issue of whether short-term forgetting reflects trace decay, interference, or both, has—over the last 40 years—been extremely difficult to settle.

Irrelevant sound effects

Students often claim that they work better to a background of their favorite music or radio program. Are they right? In 1976, Colle and Welsh showed that STM for sequences of visually presented digits was impaired when participants were required to ignore speech in an unfamiliar foreign language. However, digit recall was not impaired when irrelevant foreign speech was replaced by unpatterned noise. Salame and Baddeley (1982) went on to look directly at the effect of meaning by comparing the influence of irrelevant spoken words and irrelevant **nonsense syllables** on digit-sequence recall. They found that both meaningful words and nonsense produced an equivalent disruptive effect, suggesting that the meaning of the irrelevant material was unimportant. Indeed, recalling digit sequences when the irrelevant sounds themselves comprised digit names (*one*, *two*) caused no more disruption than when the same phonemes were presented but in a different order as nondigit words (e.g. *tun*, *woo*).

Both they and Colle (1980) suggested that the irrelevant speech effect might be seen as the memory equivalent to the masking of auditory speech perception by irrelevant sound. Perhaps the irrelevant spoken item gains access to the phonological store, and adds noise to the memory trace? However, white noise disrupts perception, but does not impair recall, whereas irrelevant speech does. Furthermore, in contrast to auditory masking, STM performance is not influenced by the intensity of the irrelevant sound (Colle, 1980). Even more problematic for the auditory masking analogy is the fact that the degree of disruption of STM is unrelated to the phonological similarity between the irrelevant sound and the items remembered. Irrelevant words that are similar in sound to the remembered items are no more disrupting than are dissimilar words (Jones & Macken, 1995; Le Compte & Shaibe, 1997).

But what about music? Salame and Baddeley (1989) found that music interfered with digit recall, finding that vocal music was more disruptive than instrumental. Jones and Macken (1993) observed that even pure tones will disrupt performance, provided they fluctuate in pitch. They proposed what they termed the *Changing State hypothesis*. This assumes that retention of serial order, whether in verbal or visual memory, can be disrupted by irrelevant stimuli providing that these fluctuate over time (Jones, Macken, & Murray, 1993). Jones (1993) relates the **irrelevant sound effect** to theories of auditory perception, presenting as an alternative to the phonological loop hypothesis their *Object-Oriented Episodic Record (O-OER) hypothesis*. This is discussed later.

The problem of serial order

It was clear by this time that the purely verbally specified phonological loop model had two major shortcomings. First, it had no adequate explanation of how serial order is stored. Given that the classic digit span task principally involves retaining serial order, this is clearly a major limitation. Second, the model has no clear specification of the crucial processes involved in *retrieval* from the phonological store. Both of these limitations demand a more detailed model, preferably computationally or mathematically simulated so that clear predictions can be made and tested. Fortunately, it has proved possible to convince a number of groups with the appropriate skills that this is a worthwhile enterprise.

A number of models based on the phonological loop have been developed, handling the question of serial order in somewhat different ways, agreeing on which issues are important but differing on how best to tackle them. The various models tend to agree in assuming both a phonological store,

KEY TERMS

Nonsense syllables: Pronounceable but meaningless consonant-vowel-consonant items designed to study learning without the complicating factor of meaning.

Irrelevant sound effect: A tendency for verbal STM to be disrupted by concurrent fluctuating sounds, including both speech and music.

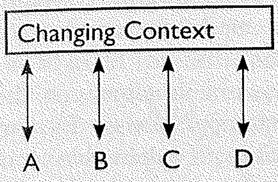
Box 2.3 Methods of storing serial order

1. Chaining

A → B → C → D

Each item is associated with the next. Recall begins with the first item (A), which evokes the second (B).

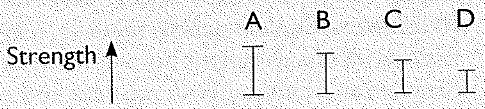
2. Context



Each item is linked to a changing context, which may be time-based. The context then acts as recall cue.

3. Primacy

Each item presented receives activation. The first receives the most, the next a little less, and so forth. Items are recalled in order of strength. Once recalled, that item is suppressed and the next strongest chosen.



and a separate mechanism for serial order, with similarity influencing retrieval from the store. Most phonological-loop-related models reject a chaining interpretation of serial order, proposing instead that order information is carried either by some form of ongoing context (Burgess & Hitch, 1999, 2006), by links to the first item as in the *Primacy Model* of Page and Norris (1998), or links to both the first and last items (Henson, 1998). Rehearsal is assumed to involve the retrieval of items from the phonological store and their subsequent re-entry as rehearsed stimuli.

Only one of these models has so far explicitly addressed the irrelevant sound effect. Page and Norris (2003) suggest that irrelevant speech adds noise to the serial order mechanism, rather than to the phonological store, the component responsible for the effects of phonological similarity (Box 2.3). This provides an explanation of why similarity between the remembered items and the irrelevant items has no effect; they influence different parts of the system and hence do not interact.

COMPETING THEORIES OF VERBAL SHORT-TERM MEMORY

We have so far focused our discussion mainly on the explanation of short-term verbal memory offered by the phonological loop hypothesis. This approach has two advantages: it provides a coherent account of a range of very robust STM phenomena, and it does so in a way that explicitly links them to those other aspects of working memory that will be discussed in the next chapter. It is important to bear in mind, however, that other ways of explaining these data have been proposed. Some of these will be described briefly before moving on to a broader discussion of working memory, and the question of why we need a working memory.

One approach to STM is that proposed by Dylan Jones and colleagues (Jones, 1993; Jones & Macken, 1993, 1995) as the Object-Oriented Episodic Record (O-OER) theory of STM. This was developed to account for the influence on STM of irrelevant sound. It is influenced by research on auditory perception and proposes that sequences of items are represented as points on a multimodal surface; however, it assumes that both auditory and visual serial recall involve the same system operating on a common representation. Recall involves retrieving the trajectory of the points representing the sequence, rather like reading off points on a graph. Irrelevant sounds might create competing trajectories, disrupting subsequent recall (Jones, 1993).

The assumption that verbal and visual STM involve the same system has not been supported by subsequent experiments (Meiser & Klauer, 1999) and is inconsistent with evidence from neuropsychological studies in which some patients are found with impaired verbal and preserved visual STM (Shallice & Warrington, 1970; Vallar & Papagno, 2002) whereas other patients show the opposite pattern (Della Sala & Logie, 2002). Furthermore, its account of memory for serial order would appear to depend on chaining, an approach that is not well supported.

Another model that has been applied to verbal STM is James Nairne's *Feature model* (Nairne, 1988, 1990), which replaces the proposed separation between LTM and STM by proposing a single memory system in which each memory item is assumed to be represented by a set of features, which are of two basic types: *modality dependent* and *modality independent*. If you read the word *HAT*, it will have both visually dependent features, such as the case in which it is printed, and visually independent features, such as its meaning. When you hear *HAT*, rather than read it, the independent features such as meaning will be the same but the dependent features will be acoustic rather than visual. Forgetting is assumed to depend on interference, with new items disrupting the features set up by earlier items, resulting in errors in recall.

The Feature model is represented by a computer program that can be used to make predictions as to the outcome of different experimental manipulations. By making various assumptions, it is possible to use the model to account for many of the results that have been used to support the phonological loop hypothesis. The phonological similarity effect is explained on the grounds that similar items have more common features, leading to a greater likelihood that a similar but incorrect item will be retrieved. Irrelevant sound is assumed to add noise to the memory trace of each individual item. Articulatory suppression is also assumed to add noise, and in addition to be attention demanding (Nairne, 1990). By making detailed assumptions about the exact proportion of modality-dependent and modality-independent features and the relative effect on these of articulatory suppression and irrelevant sound, the feature model is able to simulate a wide range of results (Neath & Surprenant, 2003), although very little justification is given for the very specific assumptions required by the various simulations.

The Feature model does have difficulty in accounting for a number of findings. It predicts that irrelevant sound will impair recall only when it occurs at the same time as the memory items are encoded. However, it disrupts recall even when it occurs *after* presentation of the memory items, even when rehearsal is prevented by suppression (Norris, Baddeley, & Page, 2004). The Feature model also has a problem explaining why the word-length effect disappears in mixed lists of long and short words. This has led to its abandonment by some of its proponents in favor of the next model to be described: the SIMPLE model (Hulme et al., 2006; Brown et al., 2007).

Brown et al. (2007) propose a very broad-ranging memory model that they call the SIMPLE (Scale Invariant Memory, Perception, and Learning) model, which they apply to both STM and LTM. It is basically a model of forgetting based on retrieval, with more distinctive items being more readily retrievable. It places emphasis on temporal discriminability but goes beyond earlier attempts to use this mechanism to explain recency effects in free recall by developing a detailed mathematical model. It is probably too early to evaluate SIMPLE, which appears to handle free recall well but appears to be less well suited to explaining serial recall (Lewandowski, Brown, Wright, & Nimmo, 2006; Nimmo & Lewandowski, 2006). As in the case of the Feature model, SIMPLE does not currently attempt to cover the executive aspects of working memory.

A further way of modeling serial order is to assume that order is maintained by a context signal. As mentioned earlier, one of these assumes a time-based context incorporating trace decay (Burgess & Hitch, 1999, 2006). This assumption is rejected by Farrell and Lewandowski (2002, 2003), who propose—in their SOB (Serial-Order-in-a-Box) model—that order is maintained using an event-based context signal, with forgetting based on interference between events.

It might seem strange that the apparently simple task of recalling a sequence of digits in the right order should prove so difficult to explain. However, as mentioned earlier, the problem of how a system like the brain that processes events in parallel can preserve serial order has challenged theorists since it was first raised by Karl Lashley (1951), over 50 years ago.

VISUO-SPATIAL SHORT-TERM MEMORY

Imagine you are in a well lit room that is suddenly plunged into darkness. Would you be able to find the door? If there was a box of matches on the desk in front of you, would you remember that it was there? These two questions concern two related but separable aspects of visual working memory, one concerned with spatial memory (*where?*) and the other with memory for objects (*what?*). The evidence suggests that you would be able to maintain a general heading towards the door for about 30 seconds (Thomson, 1983). Your memory for precise location within reach declines rather more rapidly (Elliot & Madalena, 1987).

Spatial short-term memory

Posner and Konick (1966) required their participants to remember where along a line a stimulus had occurred. Retention was good after an unfilled delay, but declined when digit processing was required in the interval, with the degree of forgetting increasing with the complexity of the digit task. As the intervening task was not spatial or visual in nature, the implication is that it interfered with capacity to rehearse or hold in mind the initial stimulus. A similar result was obtained by Dale (1973) using a task involving retention of the location of a black dot in a two-dimensional white field.

Object memory

Irwin and Andrews (1996) presented their participants with an array of different colored letters. After a brief delay, an asterisk appeared in the location previously occupied by one of the letters. This was a cue to recall the letter, the color, or both. Participants were just as good at reporting both letter and color as they were at reporting just one of these features. Performance was highly accurate up to four items, after which it declined with number of letters. It was of course possible that participants were relying on the phonological loop for the letter names, complicating interpretation.

This was avoided in a series of studies by Vogel, Woodman, and Luck (2001), who used as stimuli bars of varying width, orientation, and other features such as texture, making accurate verbal coding in the brief presentation phase virtually impossible. They again found a limit of four objects. Interestingly, it did not appear to matter how many features each object comprised. In one condition, for example, the array contained a total of sixteen features (four objects each with four features), but all four complex objects appeared to be encoded effectively. Vogel et al. concluded that the system was limited on the basis of the number of *objects* but that objects could vary in complexity without affecting performance (Figure 2.6). In one study, articulatory suppression was used to prevent verbal coding; this had no effect on performance, confirming their suggestion that their task does not depend on verbalization. They also found no evidence of forgetting over a brief unfilled delay.

Woodman and Luck (2004) interposed a visual search test in which the participants scanned an array looking for visually specified targets in the interval between presentation and test of sets of colored shapes. They found that memory for spatial location was impaired by scanning, but that object memory, as reflected in memory for the color of the stimulus, was not. Oh and Kim (2004) studied the effect of memory on the *speed* of visual search, finding that the requirement to concurrently remember a set of *objects* had no impact on visual search rate, whereas a concurrent *spatial* retention task slowed down the rate at which people could search far a specified target.

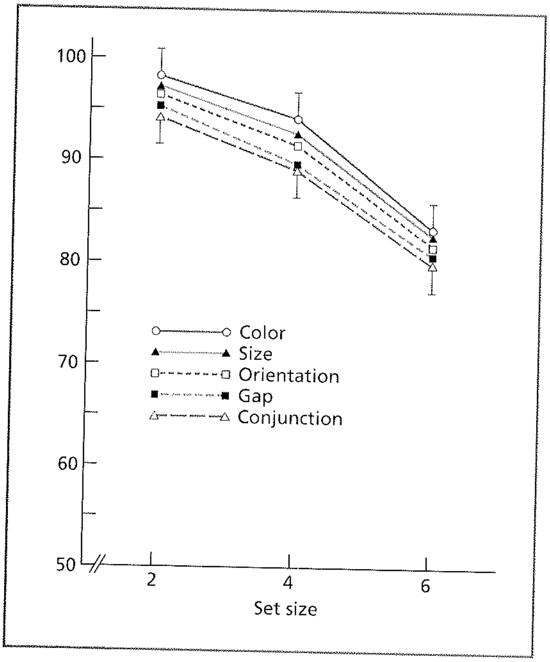


Figure 2.6 Visual recognition performance as a function of number of objects presented and number of features per object. Performance is very sensitive to number of objects but not to number of features each object comprises. Data from Vogel et al. (2001).

It therefore appears to be the case that we can remember up to four objects, even when each comprises a number of different features, and that this does not decline over a period of a few seconds, even when interpolated activity is involved. However, memory for *spatial location* appears to be more vulnerable, and to interact with other spatial activities. It seems possible that this difference might reflect the need to maintain a *spatial framework* if a precise response must be made, a capacity that might be disrupted by other spatially related activities, including the kind of eye movements that are likely to be involved in visual search.

The visual–spatial distinction

We have made a distinction between spatial STM—remembering *where*, and object memory—remembering *what*. In practice, these two systems work together but tasks have been developed that particularly emphasize one or other of these two forms of visuo-spatial memory. A classic *spatial* task is the block tapping test in which the participant is faced with an array of nine blocks (Figure 2.7). The experimenter taps a number of blocks in sequence and the participant attempts to imitate this, with the length of sequence increasing until performance breaks down. This is known as *Corsi span*, after the Canadian neuropsychologist who invented it, and is typically around five blocks, usually about two items below digit span.

Visual span can be measured using a series of matrix patterns in which half the cells are

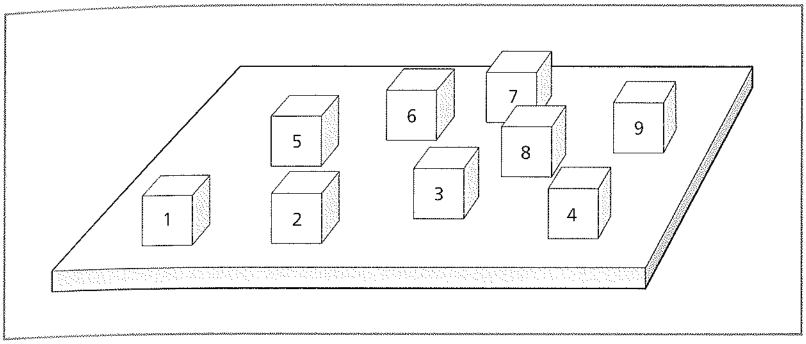


Figure 2.7 The Corsi test of visuo-spatial memory span. The experimenter taps a sequence of blocks and the participant seated opposite attempts to imitate. The numbers are there to help the experimenter.

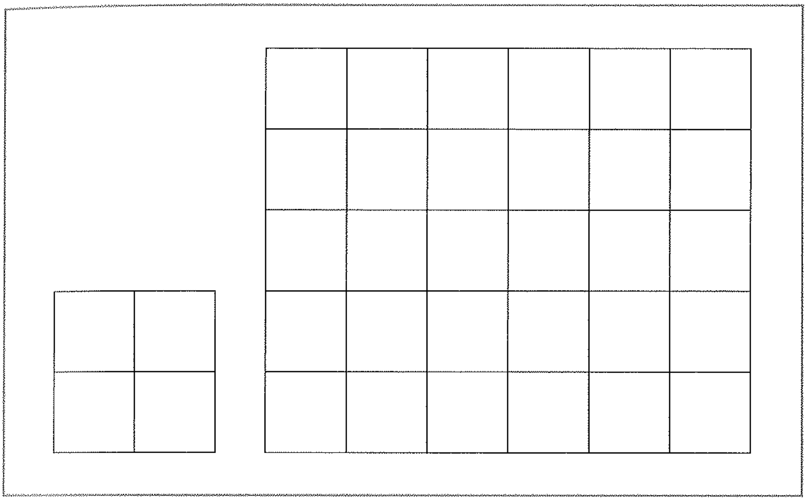


Figure 2.8 Visual pattern span. Participants are shown and then required to recall a series of matrices with half the cells filled. The test begins with a simple 2 × 2 matrix (left) and increases in size up to the 5 × 6 matrix (right). From Della Sala et al. (1999). Copyright © Elsevier. Reproduced with permission.

filled and half left blank (Figure 2.8). The participant is shown a pattern and asked to reproduce it by marking the filled cells in an empty matrix. Testing starts with a simple 2 × 2 pattern, and the number of cells in the matrix is gradually increased to a point at which performance breaks down, usually around the point at which the matrix reaches around 16 cells.

Evidence for the distinction between these two measures of spatial and visual span comes from studies in which a potentially interfering activity is inserted between presentation and test. When this involves spatial processing, such as sequentially tapping a series of keys, Corsi span is reduced; pattern span is more disrupted by a visual processing task such as viewing shapes (Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999).

Visual STM is not of course limited to remembering patterns, but also involves shapes and colors. This is shown particularly clearly in a series of studies by Klauer and Zhao (2004) in which they contrast a spatial task that involves remembering the location of a white dot on a black background, and a visual task involving memory for Chinese ideographs. In each case, the stimulus is presented and followed by a 10-second retention interval, after which participants must choose which of eight test items has just been presented. During the 10-second delay, participants perform either a spatial or a visual task. In the spatial task, 12 asterisks are presented, with 11 moving randomly and the twelfth stationary; the task is to identify the stationary item. The visual interfering task involves processing a

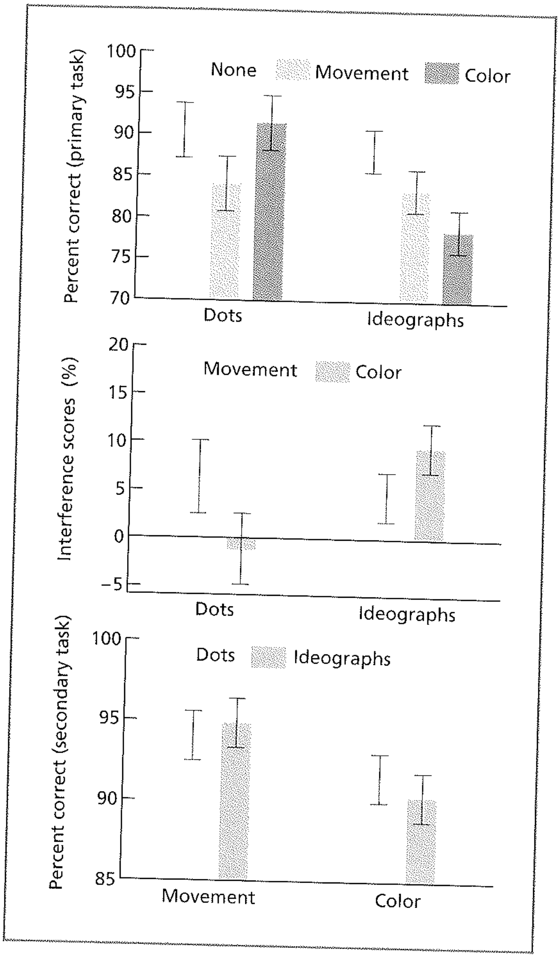


Figure 2.9 Memory for dot location and for Chinese ideographs. Only movement disrupts spatial memory, whereas pattern memory is more disrupted by intervening colors. Data from Klauer and Zhao (2004).

series of colors, seven of which are variants of one color, perhaps red, whereas one, the target, is in the blue range. As shown in Figure 2.9, the spatial location of dots was disrupted by movement but not color, whereas ideographs showed the opposite effect.

Neuropsychological approaches to the study of short-term memory

In 1966, Brenda Milner reported the case of a young man, HM, who had had the misfortune to suffer from intractable epilepsy. This can be the result of scar tissue within the brain that leads to excessive electrical activity, ending in epileptic seizures. It is often the case that surgical removal of the scar tissue greatly alleviates the seizures. HM had such an operation and his seizures were reduced, but unfortunately the operation left him densely amnesic. The reason for this was that surgical lesions had been made on both sides of the brain, in the region of the hippocampus, which we now know plays a crucial role in long-term episodic memory.

The case of HM was important practically, because it highlighted the need to take careful account of the function of the area of brain being removed and—theoretically—because of the light that it cast upon the nature of LTM. HM was impaired in his capacity to learn new material, whether visual or verbal, and to update his knowledge of the world, a classic case of the amnesic syndrome that is discussed in Chapter 11. There were aspects of memory, however, that were preserved in HM. He could remember events that had happened before his operation and he could learn certain tasks, such as those involving motor skill.

HM also had preserved digit span, suggesting that STM and LTM performance might depend on different memory systems reflecting different areas of the brain. Baddeley and Warrington (1970) therefore explored the proposed STM–LTM distinction in more detail, using a group of carefully selected patients who were densely amnesic but with otherwise normal intellectual skills, and examining their performance on a range of STM and LTM tasks. Like HM, they proved to have preserved digit span. They also showed normal performance on the Peterson task, and preserved recency in free recall, while showing a marked deficit in performance on primacy and earlier items in the free recall list. These results are discussed further in Chapter 11.

Deficits in verbal short-term memory

At the same time, Tim Shallice and Elizabeth Warrington (1970) were studying a patient with the opposite pattern of memory problems.

Patient KF had a digit span of only two items, grossly impaired Peterson task performance, and showed little recency in free recall. Other patients were subsequently identified who showed an equivalent pattern (Vallar & Shallice, 1990). These two types of patient—amnesic patients with normal performance on STM and impaired LTM, and patients with impaired STM and preserved LTM—offered what is known as a **double-dissociation**. This provides strong, although not perfect, evidence for the need to assume at least two separate systems or processes. For example, if only the amnesic patients had been studied, it could be argued that their preserved capacities such as recency and digit span simply reflected easier tasks than those that were impaired. The fact that other patients show exactly the opposite rules out this interpretation.

Shallice and Warrington's patient proved not to have a general deficit in STM, but rather a specific phonological STM deficit. Consequently, his performance was much better when his digit span was tested using visual presentation, consistent with his preserved visual memory as tested on the **Corsi block tapping** test. A similar pattern was shown by patient PV (Basso, Spinnler, Vallar, & Zanobia, 1982; Vallar & Baddeley, 1987), who developed a very pure and specific deficit in phonological STM following a stroke. Her intellect and language were otherwise unimpaired, but she had a digit span of two and failed to show either a phonological similarity or word length effect in verbal STM. As is characteristic of such patients, she showed a grossly reduced recency effect in immediate verbal free recall. She did however show normal long-term recency. This was tested using a task involving the solution of a series of anagram puzzles, followed by an unexpected request for recall. Both PV and control patients showed a clear recency effect with better recall of later solutions, even though recall was unexpected and had been followed by the need to tackle later anagrams. This pattern suggests that it is not PV's capacity to use a recency strategy that is impaired, but rather her capacity to use this to boost immediate verbal memory, which presumably relies on a phonological or verbal/lexical code.

Deficits in visuo-spatial short-term memory

Whereas some patients such as KF and PV have a deficit that is limited to verbal STM, other patients show the opposite pattern with normal verbal STM and impaired performance on either visual or spatial STM measures. One such patient, LH, had suffered a head injury in a traffic accident and was grossly impaired in his capacity to remember colors or shapes. However, he had excellent memory for spatial information such as locations and routes (Farah, Hammond, Levine, & Calvanio, 1988). Another patient, LE, suffered brain damage as a result of lupus erythematosus. She also had excellent spatial memory and was well able to drive an unfamiliar route between her home and the laboratory where her cognitive skills were tested. However, she did have impaired visual memory coupled with a grossly impaired capacity to draw from memory (Wilson, Baddeley, & Young, 1999). She was a talented sculptor, who found that she had lost her capacity to visualize. She could not remember what her earlier sculptures looked like and dramatically changed her style (Box 2.4).

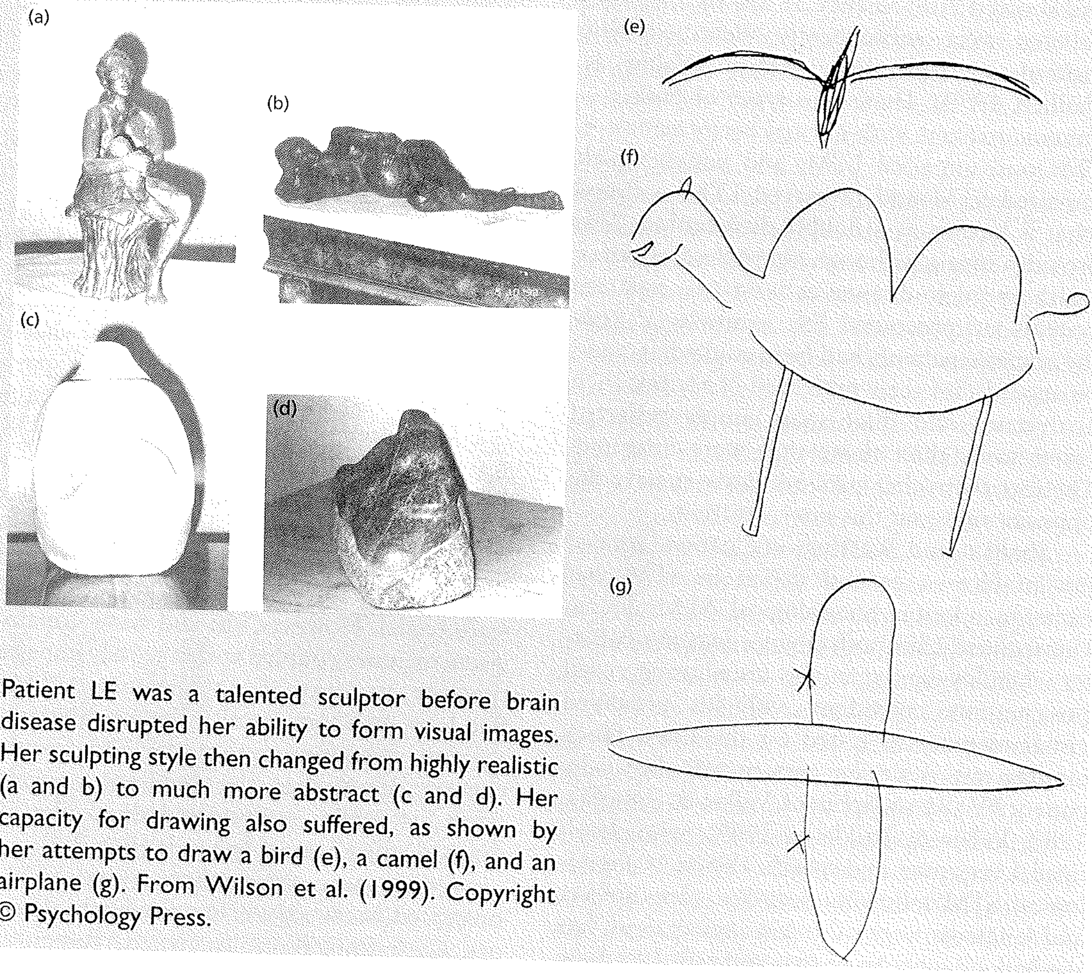
Other cases occur whose visual STM is preserved, but who have impaired spatial memory. Carlesimo, Perri, Turriziani, Tomaiuolo, and Caltagirone (2001) describe patient MV, who suffered damage to the right frontal lobe following a stroke, whose visual memory

KEY TERMS

Double-dissociation: A term particularly used in neuropsychology when two patient groups show opposite patterns of deficit, e.g. normal STM and impaired LTM, versus normal LTM and impaired STM.

Corsi block tapping: Visuo-spatial counterpart to digit span involving an array of blocks that the tester taps in a sequence and the patient attempts to copy.

Box 2.4 Patient LE



Patient LE was a talented sculptor before brain disease disrupted her ability to form visual images. Her sculpting style then changed from highly realistic (a and b) to much more abstract (c and d). Her capacity for drawing also suffered, as shown by her attempts to draw a bird (e), a camel (f), and an airplane (g). From Wilson et al. (1999). Copyright © Psychology Press.

performance was normal, but who was very impaired on the Corsi block tapping span and on a task requiring STM for imaging a path through a matrix. Luzatti, Vecchi, Agazzi, Cesa-Bianchi, and Vergani (1998) report a similar case in which progressive deterioration of the right hemisphere led to spatial memory deficits on tasks such as describing the location of landmarks in her home town, while having a good memory for colors and shapes.

You might have noticed that the deficits shown by patients with **visuo-spatial STM** problems tend to go beyond the simple storage of visual and spatial stimuli, involving more complex tasks such as creating and manipulating mental images and using these in complex tasks, such as sculpting and spatial orientation. They have, in other words, led to deficits in both STM and working memory, the topic of the next chapter.

KEY TERM

Visuo-spatial STM: Retention of visual and/or spatial information over brief periods of time.

SUMMARY

The term short-term memory (STM) refers to the storage of small amounts of information over brief periods of time. It is distinguished from working memory (WM), which is assumed to combine storage and processing and to serve as a mental workspace for performing complex tasks. The study of STM began with the digit span task used as a measure of mental capacity. Span is usually around six or seven digits, but is likely to be shorter for words and even shorter for nonsense syllables or words in a foreign language. Performance is limited by number of chunks rather than items, with different types of material being more or less chunkable.

In the late 1950s, the idea of a unitary general memory system was challenged, largely based on two experimental paradigms. The Peterson short-term forgetting effect demonstrated that small amounts of information would be forgotten within seconds if rehearsal was prevented, and was initially interpreted in terms of trace decay. It was later shown that the very first item presented showed little forgetting, suggesting an interpretation in terms of interference from prior items rather than decay. This indicates the need for a theory that incorporates retrieval.

The second influential paradigm was free recall, where the recency effect is typically lost within seconds if rehearsal is prevented, and is resistant to the many long-term memory phenomena that influenced earlier items. The existence of long-term recency suggests that it reflects a particular type of retrieval strategy that can be applied across a range of different memory systems.

Subsequent work has moved on to the study of the characteristics of various STM systems. Verbal STM has been shown to be influenced by phonological similarity and by the length of the words being retained. The phonological loop hypothesis attempts to explain this within a broader working memory framework by assuming a temporary store and an articulatory rehearsal process that can be interrupted by articulatory suppression. A number of models have been proposed to explain this overall pattern of results. Some are based on the phonological loop model, others include the Feature hypothesis, which is principally influenced by models of LTM; the Object-Oriented Episodic Record model, influenced mainly by irrelevant sound effects; and the Scale Invariant Memory, Perception, and Learning model (SIMPLE), which is strongly influenced by the need to give an account of recency.

Visual STM can be divided into visual and spatial memory. Memory for spatial location appears to show forgetting over a period of seconds, while memory for visual objects does not. We appear able to retain up to four objects, with performance declining beyond this point. Somewhat surprisingly, the number of features an object comprises does not seem to be obviously limited. Visual and spatial components have been proposed as part of the visuo-spatial sketchpad, a component of working memory that is a counterpart of the phonological loop.

Neuropsychological evidence supports the behavioral distinctions proposed. A number of cases have revealed the potential separation between verbal STM and both LTM and visual STM, with the pattern of deficits being broadly consistent with the phonological loop hypothesis.

Patients with visuo-spatial STM deficits have also been reported, some who appear to have disruption of the visual store, reflected in impaired pattern span, whereas others show impaired spatial STM as reflected in the Corsi block tapping task. The lesion sites leading to these neuropsychological deficits are broadly consistent with evidence from neuroimaging that are discussed in Chapter 3.

FURTHER READING

- **Logie, R. H.** (1995). *Visuo-spatial working memory*. Hove, UK: Psychology Press. An account of visual STM from a multicomponent working memory perspective.

- **Luck, S. J., & Vogel, E. K.** (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281. An important paper that forms a link between the study of visual attention and visual STM.

- **Melton, A. W.** (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 2, 1–21. A classic paper presenting an interpretation of STM in terms of stimulus–response interference theory.

- **Vallar, G.** (2006). Memory systems: The case of phonological short-term memory. A festschrift for *Cognitive Neuropsychology. Cognitive Neuropsychology*, *23*, 135–155. An explanation and account of phonological loop from the viewpoint of neuropsychology.

- **Waugh, N. C., & Norman, D. A.** (1965). Primary memory. *Psychological Review*, *72*, 89–104. Another classic paper presenting an information-processing alternative to interference theory.